Documentation projet InvestQ

# Liste de ressources utilisées

* [Exemple de projet d’apprentissage par renforcement](https://medium.datadriveninvestor.com/how-to-use-reinforcement-learning-for-profitable-investing-ebe8662f13cc)
* [Documentation de l’API Alpaca](https://docs.alpaca.markets/reference/stockbars-1)
* [Article 2018](https://arxiv.org/pdf/1811.07522)

# Notions et vocabulaires

## Titre

Un titre(stock), également connue sous le nom « equity », est un titre financier qui représente la propriété d'une fraction de la société émettrice. Les unités d'actions sont appelées parts (shares), qui donnent à leur propriétaire droit à une proportion des actifs (assets) et des bénéfices de la société correspondant à la quantité d'actions qu'il possède.

Les actions sont principalement achetées et vendues sur les bourses et constituent la base des portefeuilles de nombreux investisseurs individuels. Les transactions boursières doivent respecter les réglementations gouvernementales destinées à protéger les investisseurs contre les pratiques frauduleuses.

## Positions

Une position est la quantité d'un titre, d'un actif ou d'une propriété possédée (ou vendue à découvert) par un individu ou une autre entité. Un trader ou un investisseur prend une position lorsqu'il effectue un achat via un ordre d'achat, indiquant une intention haussière ; ou lorsqu'il vend des titres à découvert avec une intention baissière.

L'ouverture d'une nouvelle position est finalement suivie à un moment donné dans le futur par la sortie ou la clôture de la position (c’est-à-dire la vente de tous les parts).

### Position longue

Une position longue signifie que l'investisseur a acheté un actif financier (comme des actions, des obligations ou des marchandises) et s'attend à ce que le prix de cet actif augmente.

L'objectif est de vendre l'actif à un prix plus élevé dans le futur pour réaliser un profit.

### Position courte

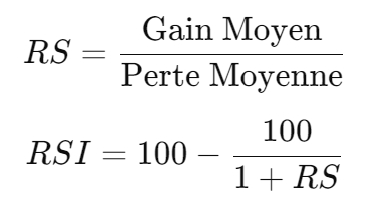
Une position courte signifie que l'investisseur a emprunté un actif financier et l'a vendu avec l'intention de le racheter plus tard à un prix inférieur.

L'objectif est de profiter de la baisse du prix de l'actif en le rachetant à un prix inférieur à celui de la vente initiale.

# Indicateurs techniques d’analyse

## L’indicateur RSI

L'indice de force relative (RSI) est un indicateur technique populaire utilisé dans l'analyse des marchés financiers. Il est conçu pour mesurer la vitesse et le changement des mouvements de prix. Le RSI oscille entre 0 et 100 et est généralement utilisé pour identifier les conditions de surachat ou de survente dans un marché.



### L’indicateur MACD

Le MACD (Moving Average Convergence Divergence) est un indicateur technique utilisé dans l'analyse technique des marchés financiers. Il est basé sur les différences entre deux moyennes mobiles exponentielles (EMA) et est souvent utilisé pour identifier les tendances et les points d'inversion potentiels.

### L’indicateur ADX

L'indice directionnel moyen (ADX) est un indicateur d'analyse technique utilisé par certains traders pour déterminer la force d'une tendance.

La tendance peut être soit à la hausse soit à la baisse, et cela est montré par deux indicateurs d'accompagnement, l'indicateur directionnel négatif (-DI) et l'indicateur directionnel positif (+DI). Par conséquent, l'ADX comprend généralement trois lignes distinctes.

2 indicateurs de performances :

* Sharpe ratio
* Cumulative returns

# Récupération des données historiques

L’historique d’Alpaca remonte jusqu’à 2016 mais avec des données manquantes. On utilisera les données depuis 2017 afin qu’elles soient plus complète.

4 types de données historiques :

* Historical bars : correspond aux candlesticks qu’on va détailler ci-dessous
* Historical auctions : …
* Historical quotes : …
* Historical trades : …

**SIP** is short for [Securities Information Processor](https://en.wikipedia.org/wiki/Securities_information_processor). All US exchanges are mandated by the regulators to report their activities (trades and quotes) to the consolidated tape. This is what we call SIP data.

**IEX** ([Investors Exchange](https://en.wikipedia.org/wiki/IEX" \t "_self)) is a single stock exchange.

On utilisera le SIP pour les données historiques et IEX pour récupérer les données en temps réel étant donné qu’on n’a pas accès aux données SIP en temps réel avec la version gratuite de l’API Alpaca.

On récupère une liste de stocks avec 2 times frame différents : L’état du stock sur une durée de 1 jour et sur 1 heure.

API limitée à 200 requêtes par minute => il faut donc ajouter un sleep pour éviter un blocage de l’API surtout pour un time frame d’une heure (beaucoup plus de données et donc d’appel à l’API)

On récupère des données sous la forme de « candlestick ». Cette forme permet de représenter l’état du marché durant l’intervalle de temps défini, en indiquant :

* Le prix d’ouverture -> la valeur de la première transaction
* Le prix de fermeture -> la valeur de la dernière transaction
* Le prix le plus élevé durant cet intervalle de temps
* Le prix le plus bas durant cet intervalle de temps

Ajouté à ça :

* Le nombre d’échange effectué
* L’horodatage indiquant le début de période spécifiée
* Le volume d’actions échangées durant la période spécifiée
* Le Volume-Weighted Average Price - VWAP : Le prix moyen pondéré par le volume pour la période spécifiée. Il est calculé en multipliant chaque prix de transaction par le volume de cette transaction, puis en divisant la somme des produits par le volume total des transactions

# Etat de l’art

Blablabla article

# Expérimentation

Lorsqu’on observe l’évolution du prix en fonction du temps pour une action, on voit clairement des tendances se dessiner. On pourrait croire naïvement qu’il suffit d’acheter lorsque l’action augmente et vendre lorsque celle-ci baisse sur des périodes suffisamment grandes. Pour se rendre compte de la difficulté du problème, on va simuler à l’aide des données historiques des stratégies simples : La première basée sur la différence de prix entre deux temps différents et une autre sur le calcul du momentum afin de lisser les variations et mieux faire ressortir les tendances. Ces deux stratégies ne sont pas officielles et sont seulement là pour montrer que ce n’est pas un problème simple.

Dans un second temps, on va utiliser une approche machine learning avec du forecasting sur des séries temporelles avec un modèle basé sur des couches LSTM puis un modèle plus complexe avec le Prophet de Méta. Finalement, on va tester des algos d’apprentissage par renforcement en utilisant des modèles comme PPO qui sont efficient sur des environnements stochastiques.

Les résultats que l’on va présenter ci-dessous se base sur l’action d’Apple entre le 1er janvier 2023 jusqu’au mois de mai 2024 (1 an et demi). Chaque valeur est éloignée d’une heure l’une de l’autre (time frame d’une heure). Pour simplifier le problème, on ne peut acheter qu’une seule action à la fois et on considère un portfolio de 10.000€ initialement. Pour les modèles de machine learning on va utiliser les données entre le 1er janvier 2017 jusqu’à 2022 pour l’entrainement et jusqu’à 2023 pour le jeu de validation, le jeu de test sera le même entre le 1er janvier 2023 jusqu’au mois de mai 2024. On notera que Apple à diviser son action par 4 au 31 Aout 2020, on va donc multiplier par 4 les prix pour ne pas perturber l’apprentissage de nos modèles.

## Approche simple

### Différence de prix

On va commencer par calculer la différence de prix entre deux temps différents que l’on va faire varier. L’idée est que si on est dans une tendance ou le prix augmente, alors le prix actuel doit être supérieur au prix précédent et donc il faut acheter. Inversement lorsqu’on a une tendance à la baisse. On va également faire varier l’intervalle de temps pour ce calcul de différence de prix. En effet, à cause de la forte volatilité des prix il pourrait être intéressant de se baser sur des temps plus larges.

Pour comparer les résultats on utilisera comme baseline, l’évolution de l’action.

Une image contenant texte, capture d’écran, Police, Tracé

Description générée automatiquement

On affiche ci-dessus les résultats pour un calcul de différente entre 1, 10, 20, 40 et 80 valeurs qui ne correspondent pas exactement à leur équivalent en heure car le marché n’est pas ouvert 24h/24.

Comme on peut l’observer, seulement la courbe verte (différence de 40) est supérieure à l’action à la fin de l’historique et le gain n’est pas significativement meilleur. On ne peut pas conclure que cette stratégie de trading est plus performante que simplement se positionner sur une action à long terme.

### Momentum du prix

Afin de faire ressortir les tendances, on va lisser le bruit en utilisant le momentum du prix.



Avec pt le prix au temps t, beta un paramètre entre 0 et 1, et mt le momentum calculé.

On va utiliser la même stratégie qu’avant pour calculer la différence de prix entre le prix actuel et la valeur précédente mais basée sur le momentum du prix. On choisit les valeurs béta suivantes [0.9, 0.95, 0.99, 0.995, 0.999]. Plus la valeur est grande est plus le prix sera lissé mais le momentum sera davantage en retard sur le prix réel.

Une image contenant texte, capture d’écran, Police, ligne

Description générée automatiquement

On obtient les meilleurs résultats pour les valeurs de béta 0.99 et 0.995. Cette stratégie semble plus efficace que la précédente mais le gain reste trop faible pour pouvoir conclure à sa réelle efficacité. Sur d’autres périodes ou actions les paramètres optimales pourraient être différents et rien n’assure d’avoir de meilleur résultat.

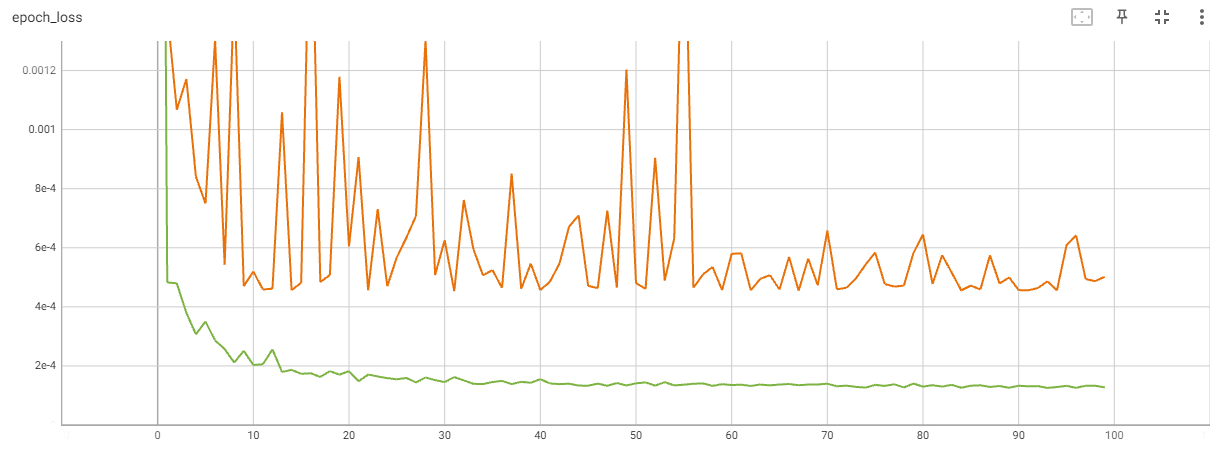
Ces 2 expérimentations démontrent bien la difficulté du problème. Effectuer du trading en se basant uniquement sur la tendance aboutit généralement à de moins bon résultats qu’une position longue sans effectuer de trading. On a pu optimiser les paramètres pour obtenir des résultats légèrement meilleurs que l’évolution de l’action parce qu’on connaissait l’évolution du prix de l’action sur toute la période de trading. Dans un cas réel ou on ne connaitrait pas l’évolution des prix futur, on ne pourra pas déterminer les meilleurs paramètres.

Les traders utilisent de multiples indicateurs plus sophistiqué, leurs expériences personnelles et les évènements qui pourrait influencer l’action en dehors de l’étude pur de l’évolution des prix.

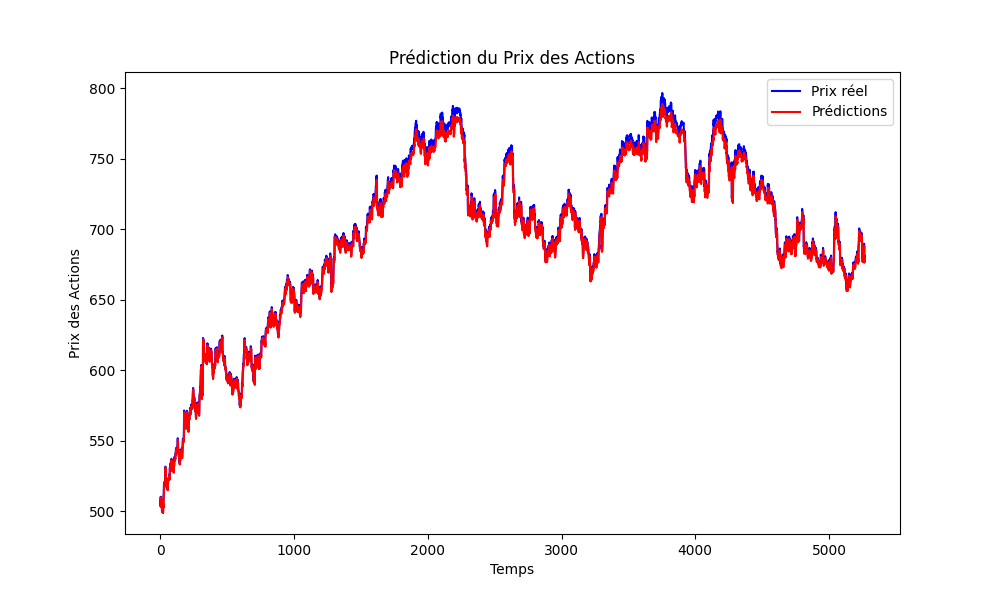
## Approche machine learning

### Approcher une série temporelle avec des réseaux de neurones récurrent LSTM

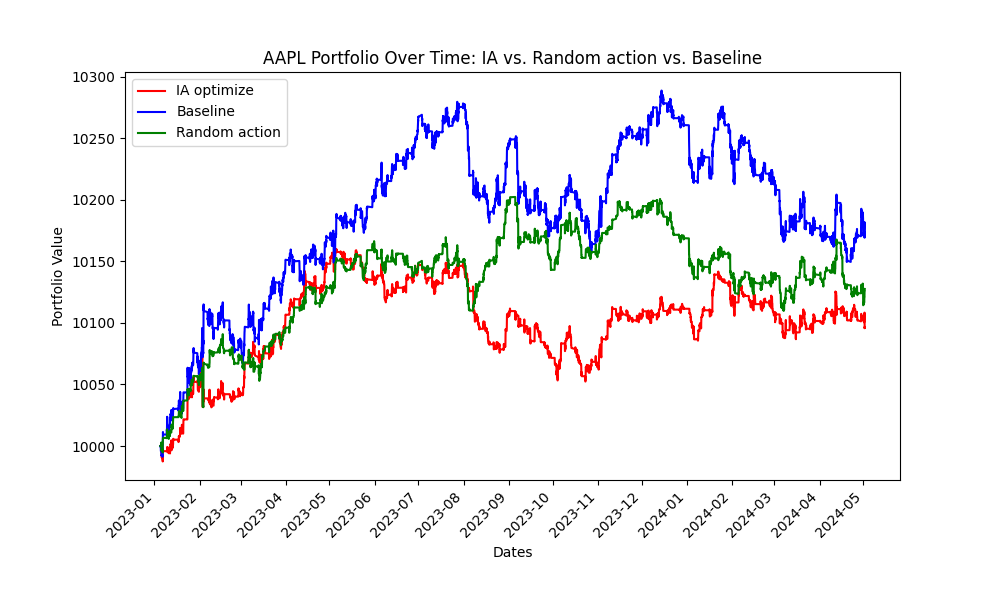
Les réseaux de neurones récurrents sont adaptés à la modélisation de série temporelle. On va essayer un modèle simple constitué de 2 couches LSTM de 128 neurones avec uniquement les prix de fermeture en entrée. La fenêtre d’observation est de 30 valeurs, on applique une normalisation standard sur nos données étant donné qu’on a des valeurs bien plus grande que 1.



Le modèle a bien convergé avec un léger overtfiiting (validation en orange > train en vert)



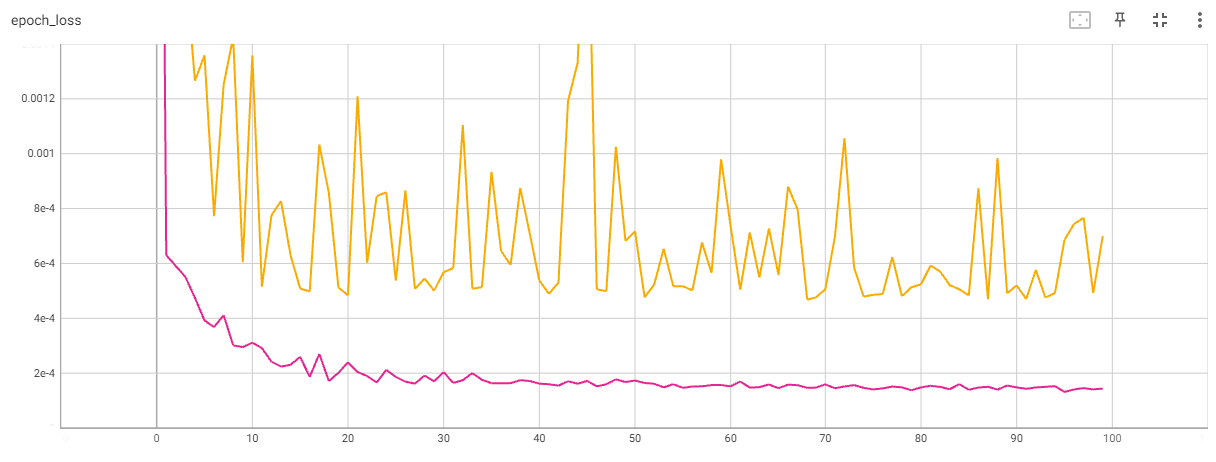
Le graph ci-dessus représente le prix réel vs le prix prédit. On prédit le prix au temps suivant à partir des 30 derniers prix réels. Les résultats semblent très bien fit mais cela ne prouve en rien que le modèle est capable de prédire le prix suivant, en effet il suffit au modèle de prédire une valeur similaire à la dernier valeur pour avoir un résultat semblable. Cela est clairement visible si on calcule la proportion de bonne prédiction à la hausse ou à la baisse des prix. C’est-à-dire si le modèle prédit un prix plus élevé qu’au temps précédent et qu’il s’avère effectivement plus élevé on considère que c’est une bonne réponse et de même pour les prédictions à la baisse. On obtient pour ce modèle 49.3% de bonne prédiction qui est totalement aléatoire car on a une chance sur deux de se tromper soit 50%.

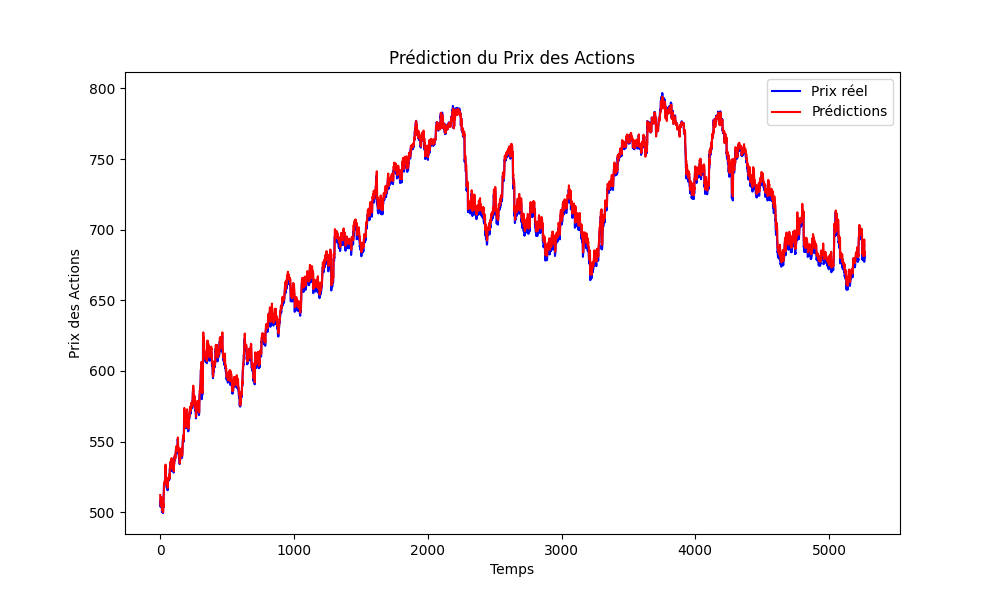


Le graphique ci-dessus représente l’évolution du portfolio de l’IA vs l’évolution de l’action et des actions aléatoires. Si on prédit que le prix va augmenter, on achète sinon on vend. Comme énoncé précédemment, même si le modèle est capable de prédire une valeur proche de la vraie valeur, il n’a aucune idée si elle sera plus grande ou plus faible. Ainsi, l’IA a des performances similaires à l’aléatoire.

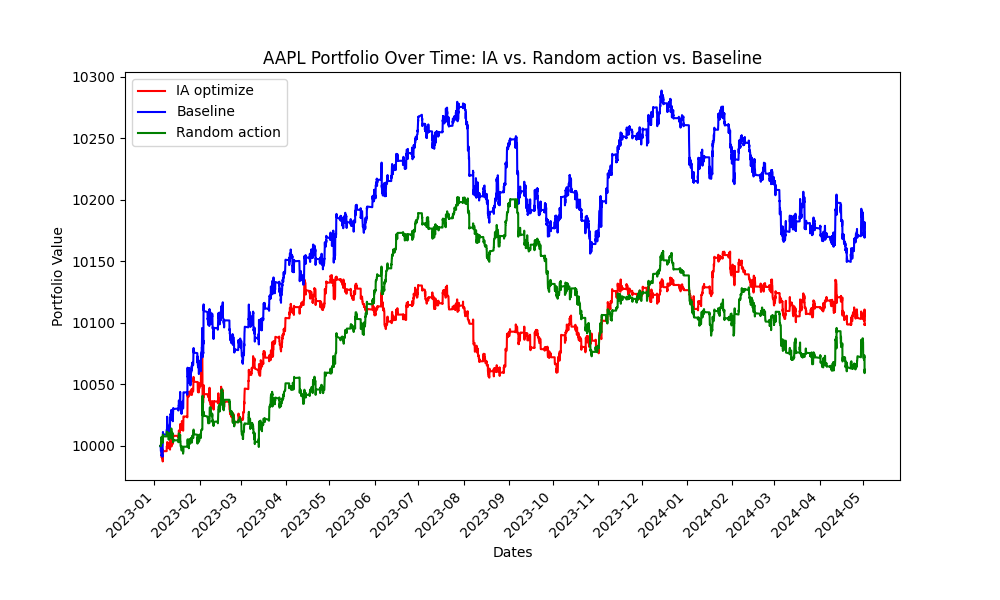
Maintenant en plus des prix de fermeture, on va ajouter le nombre de transaction et le nombre total de volume d’action échangé. On veut tester si ces features peuvent aider notre modèle à comprendre l’évolution des prix.

On représente ci-dessous les mêmes types de graphique :





Avec 48.6% de bonne prédiction de tendance.

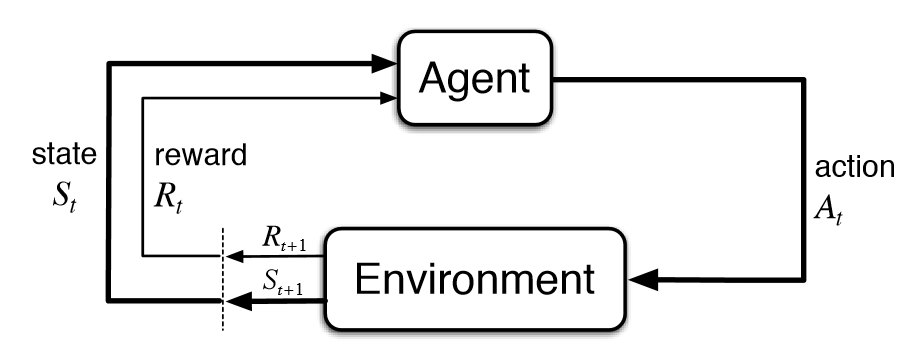


Même constat que précédemment => ces 2 features supplémentaire n’ont pas permis de mieux comprendre l’évolution des prix.

### Apprentissage par renforcement

Avec les réseaux récurrents on a essayé de prédire l’évolution des prix pour prendre une décision sur l’achat ou la vente de l’actif. Cette fois-ci avec des modèles d’apprentissage par renforcement, on va demander à notre modèle de directement nous donner l’action à effectuer, acheter, vendre ou ne rien faire.

Le schéma ci-dessous représente le principe de l’apprentissage par renforcement :

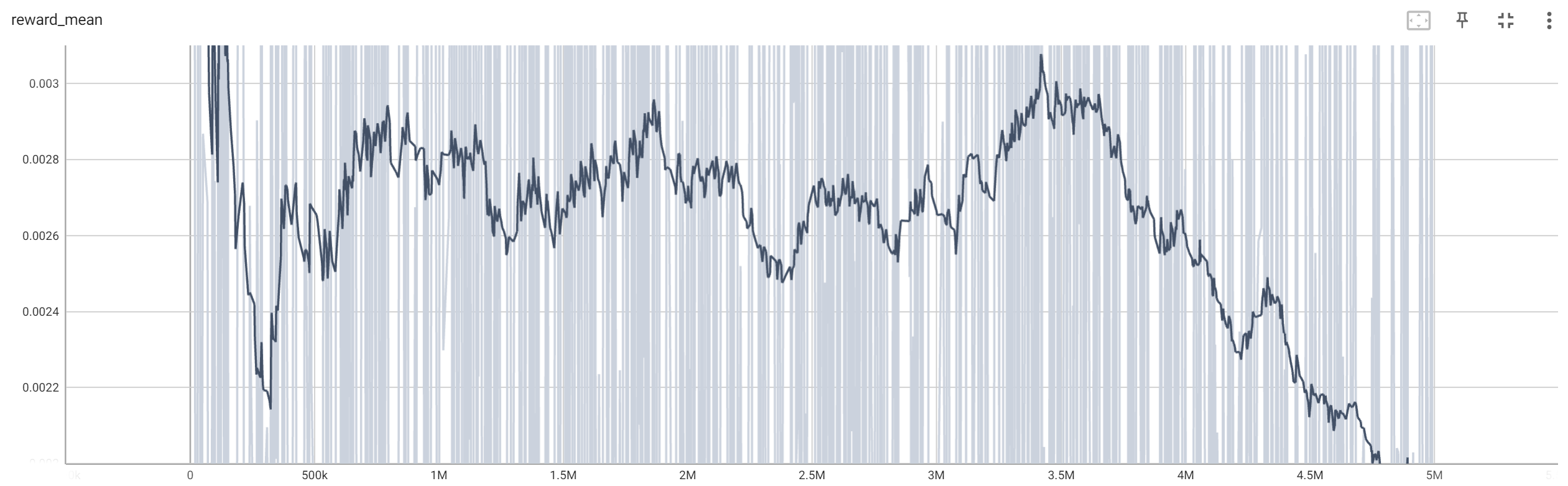


Le modèle ou agent, effectue une action sur l’environnement à partir de l’observation de celui-ci. Au temps suivant, l’environnement renvoie la nouvelle observation et la récompense associé à l’action prise. A partir de la nouvelle observation, l’agent effectue une nouvelle action et ainsi de suite. La récompense permet à l’agent d’avoir un retour sur l’action qu’il a effectué, ce qui va lui permettre d’apprendre. En effet, l’agent va chercher à maximiser la somme cumulée de toutes les récompenses.

Concernant notre problématique, on souhaite maximiser la valeur de notre portfolio au cours du temps. Naturellement la récompense sera la valeur du changement du portfolio. L’environnement sera l’évolution des prix. L’observation est plus complexe, on peut récupérer uniquement le dernier prix et ajouter des indicateurs calculés sur une certaine période passée afin de donner une idée au modèle de la tendance des prix par exemple ou simplement l’historique des prix comme pour le modèle précédent LSTM. On peut également y ajouter le nombre d’action en possession, la valeur du portfolio actuelle etc.

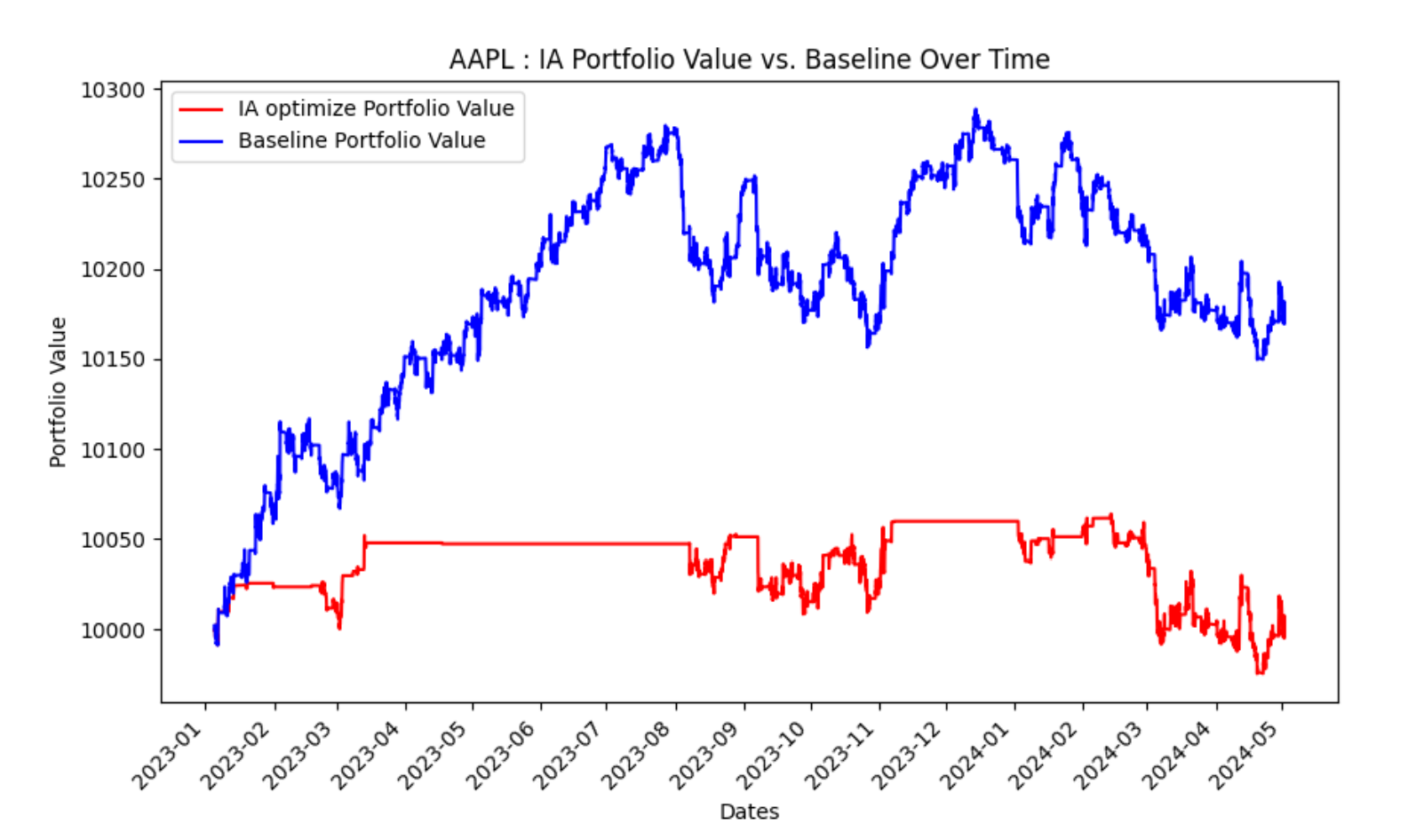
Comme précédemment, afin de simplifier le problème on permet au modèle de posséder qu’une unité d’action à la fois. Le réseau de neurones est composé de 2 couches denses de 64 unités.

On choisit en input l’historique des 30 valeurs de prix précédent et la position actuelle, c’est-à-dire si l’agent est en possession ou non d’une action. On entraine notre modèle sur 5 millions de step, c’est-à-dire 5 millions d’actions effectuées sur le jeu d’entrainement.



On représente l’évolution de la reward moyenne sur notre rollout de 2048 steps. Le rollout est la taille des données récoltés avant d’effectuer une itération d’apprentissage. Les résultats étant très instable, on adoucie la courbe afin de voir apparaitre la tendance.

La reward moyenne diminue alors qu’on s’attend à ce qu’elle augmente pour maximiser les gains. On vérifie sur le graphique suivant l’évolution du portfolio sur le jeu de test en appliquant notre modèle entrainé.

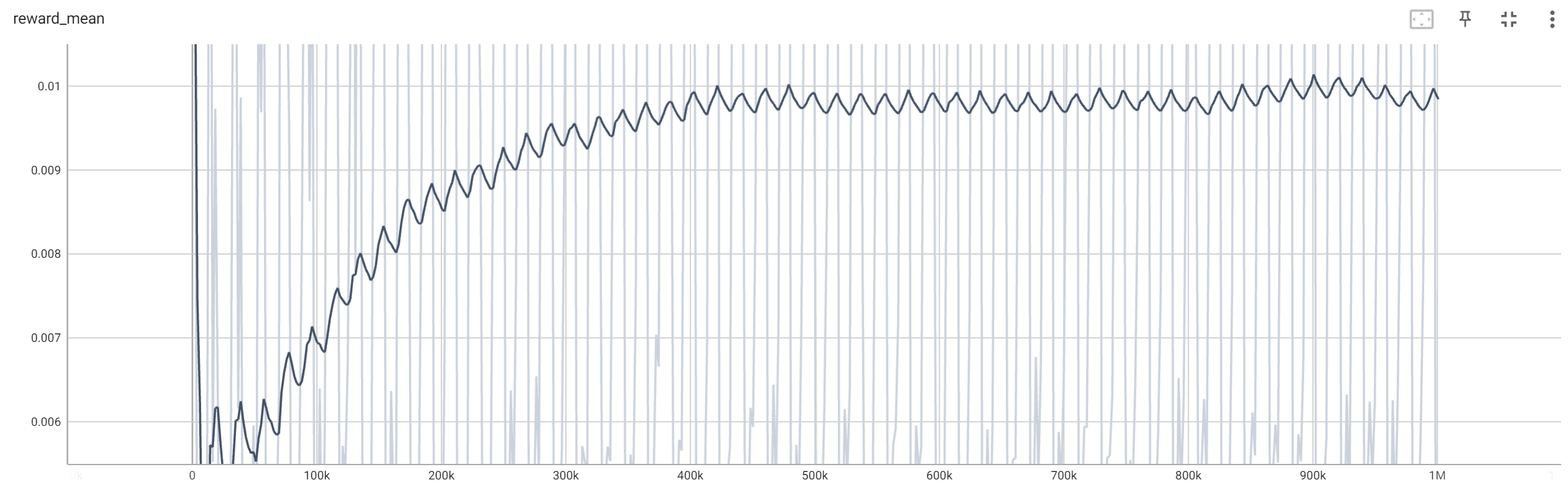


Le modèle a appris à minimiser le gain ce qui est l’inverse de ce que l’on souhaite. Effectivement la courbe en rouge représente les performances de notre modèle et en bleu la simulation d’une position longue sur toute la durée du test. La raison reste pour le moment inconnue.

Notre deuxième tentative aura une approche plus humaine. On va donner à notre modèle la dernière valeur de l’action, la position actuelle et les indicateurs suivants :

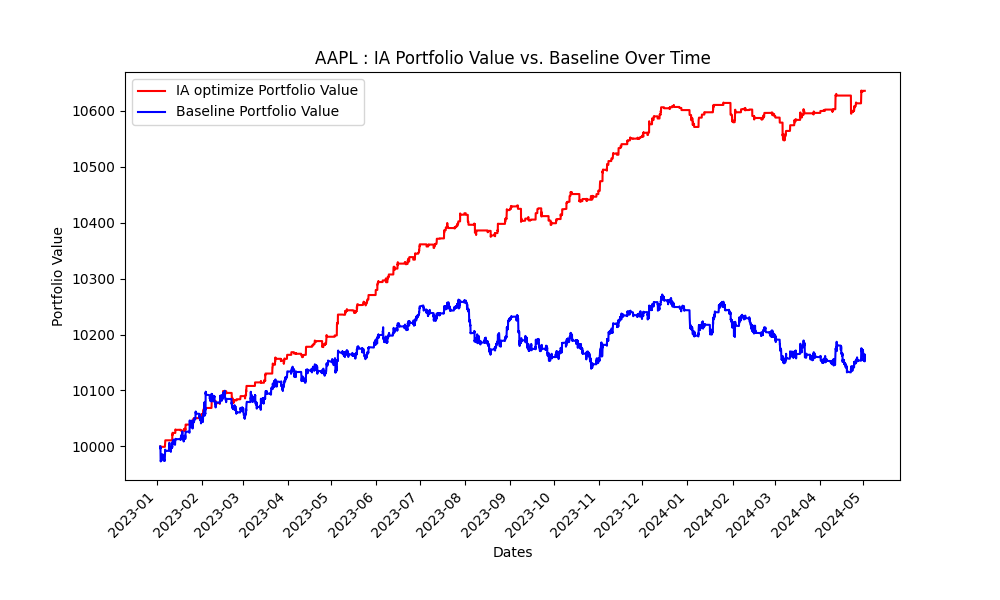
* RSI
* ADX
* MACD

Tous calculés sur un timeframe d’un jour et d’une heure. Par exemple RSI sera calculé sur une période de 14 jours mais aussi de 14 heures.



Cette fois-ci on a bien une reward moyenne qui augmente sur un total de 1 million de step.

On vérifie à nouveau en effectuant une simulation sur le jeu de test



Notre modèle a d’excellente performance et dépasse de loin notre baseline qui est équivalent à l’évolution du prix de l’action. L’entrainement a été lancé une deuxième fois pour s’assurer que ce n’était pas un coup de chance. Attention, il faudra attendre le test de notre modèle sur la bourse en temps réel avant de confirmer les performances de notre modèle.

On pourrait continuer d’expérimenter d’autres indicateurs ou de fine tune notre modèle pour gagner encore plus en performance mais dans la prochaine partie on va plutôt se concentrer sur l’investissement de plusieurs types d’action en simultané.

### Gestion de plusieurs actions

Afin de diversifier les investissements financiers il est important de ne pas tout miser sur une action mais de répartir son capital sur un ensemble d’actifs choisis judicieusement. Pour le moment on va se baser sur les 30 différentes actions de l’indice DJIA, on pourrait envisager comme amélioration future un modèle qui choisirait les actions les plus performantes du moment.